Additional File 1: Formal description of WORMSIM

Index

1. *This document*

This document provides a description of the WORMSIM model structure and default parameter quantification as used in the simulations for the paper *Feasibility of controlling hookworm infection through preventive chemotherapy: a simulation study using the individual-based WORMSIM modelling framework* by Coffeng *et al* (Parasites and Vectors 2015). Given the many similarities between WORMSIM and ONCHOSIM, this document's contents are adapted from a previously published formal description of ONCHOSIM [1].

2. *Introduction*

WORMSIM is a generalised framework for modelling transmission and control of helminth infections in humans. It is based on previous individual-based models for onchocerciasis (ONCHOSIM), schistosomiasis (SCHISTOSIM), and lymphatic filariasis (LYMFASIM) [2–4]. WORSIM simulates the life histories of individual helminths and their transmission from person to person mediated by either a cloud of vectors or an environmental reservoir. In addition, WORMSIM can be used to evaluate the effects of different control strategies, such as vector control and chemotherapy. WORMSIM combines two simulation techniques; *stochastic microsimulation* is used to calculate the life events of individual persons and their inhabitant parasites, while the dynamics of infective material in the cloud (i.e. the vector population or environmental reservoir) is simulated *deterministically*.

The version of WORMSIM used in this study (v2.58Ap9) is originally based on the C++ code of ONCHOSIM, but has been redesigned and extended using object-oriented principles and has been programmed in Java. Individual people and mature worms are modelled as distinct objects. WORMSIM is event-driven, which means that time progresses as a result of events (although monthly events are used for most processes). The main advantages of the implementation in Java are improved code quality and therefore easier maintenance and extension. Model input parameters are specified in a structured XML-file, which is automatically validated using an XML Schema before the start of a set of simulations.

The WORMSIM framework is very flexible in that it allows the user to choose probability distributions for stochastic processes (Table A1-1) and functional relationships for deterministic processes (Table A1-2), and to change the associated parameter values. Table A1-3 provides an overview of the probability distributions, functional relationships, and parameter values used in this study.

In section 3, we describe the general structure of the modelling framework. In footnotes we highlight details and alternative options that are not evident from the mathematical descriptions. In sections 0, we provide instructions for installing and running WORMSIM. In sections 5 and 6, we present annotated WORMSIM input and output files, respectively.

Table A1-1. Probability distributions available in WORMSIM. Reference numbers (rightmost column) are used in the WORMSIM input file to define probability distributions for stochastic processes. Within WORMSIM, stochastic processes are pre-defined to follow either a continuous or discrete distribution, so each type has its own list of reference numbers.

Probability distribution	Parameters in input file	Domain	Probability density function	Reference number
Continuous distributions				
Constant (real)	μ	μ	$f(x) = 1$	0
Uniform	μ, p_1	$(p_1, 2\mu - p_1)$	$f(x) = (2\mu)^{-1}$	1
Exponential	μ	$\mathbb{R}^+$	$f(x) = \mu^{-1}e^{-x/\mu}$	2
Weibull	μ, p_1	$\mathbb{R}^+$	$f(x) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^\alpha}$, where $\alpha = p_1$ and $\beta = \mu/\Gamma(1 + 1/\alpha)$	3
Gamma	μ, p_1	$\mathbb{R}^+$	$f(x) = \frac{x^{k-1}e^{-x/\theta}}{\Gamma(k)\theta^k}$, where $k = p_1$ and $\theta = \mu/p_1$	4
Beta (optionally scaled)	μ, p_1, p_2	$(0, p_2)$	$f(x_1) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_{trans}^{\alpha-1}(1 - x_{trans})^{\beta-1}$, where $x_{trans} = x/p_2$, $\alpha = p_1$, and $\beta = p_1\left(\frac{p_2}{\mu} - 1\right)$	5
Normal	μ, p_1	$\mathbb{R}$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}$, where $\sigma = p_1$	6
Log-normal	μ, p_1	$\mathbb{R}^+$	$f(x) = \frac{1}{x\sqrt{2\pi\sigma_{ln}^2}}e^{-(\ln(x)-\mu_{ln})^2/2\sigma_{ln}^2}$, where $\mu_{ln} = \ln\left(\frac{\mu^2}{\sqrt{p_1^2+\mu^2}}\right)$, and $\sigma_{ln}^2 = \ln\left(\frac{p_1^2+\mu^2}{\mu^2}\right)$; i.e. μ and p_1 are defined on the positive real plane $\mathbb{R}^+$.	7
Discrete distributions				
Constant (integer)	μ	$\|\mu\|$	$f(x) = 1$	0
Bernoulli*	μ	$\mathbb{Z}_1$	$f(x) = \mu^x(1 - \mu)^{1-x}$	1
Uniform*	μ	$\mathbb{Z}_{j+1}$	$f(x) = \frac{1}{j+1}$, where $j = \lfloor(2m + 0.5)\rfloor$ and $\lfloor y\rfloor$ is the largest integer not larger than y	2
Binomial*	μ, p_1	$\mathbb{Z}_{p_1}$	$f(x) = \binom{n}{x}\mu^x(1 - \mu)^{n-x}$, where $n = p_1$	3
Geometric*	μ	$\mathbb{Z}^*$	$f(x) = (1 - p)^x p$, where $p = \frac{1}{\mu+1}$	4
Poisson	μ	$\mathbb{Z}^*$	$f(x) = \frac{e^{-\mu}\mu^x}{x!}$	5
Negative binomial	μ, p_1	$\mathbb{Z}^*$	$f(x) = \left(\frac{k}{k+\mu}\right)^k \frac{\Gamma(k+x)}{x!\Gamma(k)}\left(\frac{\mu}{k+\mu}\right)^x$, where $k = p_1$	6

* This distribution was available in the original ONCHOSIM model, but is still due to be implemented in WORMSIM. The Bernoulli distribution is used in parts of the model, but is hardcoded in these cases. The geometric distribution can be simulated by means of a negative binomial distribution with $p_1 = 1$. In a future update of WORMSIM, all listed distributions will be (re-)implemented using the Apache Commons Math library.

Table A1-2. Functional relationships available in WORMSIM. Reference numbers (rightmost column) are used in the WORMSIM input file to define functional relationships for deterministic processes.

Functional relationship	Parameters	Formula	Parameter constraints and function characteristics	Reference number
Constant	a	$f(x) = a$	$a \geq 0$	0
Linear	a, b, c	$f(x) = ax + b$ if $ax + b < c$, $f(x) = c$ otherwise	$a \geq 0$; $b \geq 0$; if $c < 0$, no maximum is considered	1
Hyperbolic saturating	a, b, c	$f(x) = c + \dfrac{ax}{1 + ax/(b - c)}$	$b \geq 0$; $c \geq 0$; $(b - c)a > 0$; $f(0) = 0$; $\lim_{x \to \infty} f(x) = b$; $f'(0) = a$	2
Exponential saturating	a, b, c	$f(x) = a(1 - e^{-bx})(1 + e^{-cx})$	$a \geq 0$; $b > 0$; $c \geq 0$; $f(0) = 0$; $\lim_{x \to \infty} f(x) = a$; $f'(0) = 2ab$	3
Sigmoidal saturating	a, b, c	$f(x) = a\left(1 - e^{-(bx)^c}\right)$	$a \geq 0$; $b > 0$; $0.1 < c < 10$; $f(0) = 0$; $\lim_{x \to \infty} f(x) = a$; $f'(0) = 0$ for $c > 1$; $f'(0) = ab$ for $c = 1$; $f'(0) = \infty$ for $c < 1$	4
Power function	a, b	$f(x) = ax^b$	$a \geq 0$; $b \geq 0$; $f(0) = 0$; $\lim_{x \to \infty} f(x) = \infty$; $f'(0) = 0$ for $b > 1$; $f'(0) = a$ for $b = 1$; $f'(0) = \infty$ for $b < 1$	5

Table A1-3. WORMSIM quantification used to simulate hookworm transmission. Given that WORMSIM is a general modelling framework that covers both STH and filariasis, certain parameters do not apply to hookworm transmission but are listed anyway for the sake of completeness (indicated where applicable). Further, within the groups of model parameters for transmission and surveys, certain parameters are strongly correlated as indicated by "*not identified*". Where this is the case, set all but one parameter to arbitrary values, and then used the one parameter to tune the model, as indicated by "*used as main parameter…*".

Parameter	Value	Source
Human demography		
Cumulative survival ($F(a)$), by age (see also Figure A1)		[5]
0	1.000	
5	0.804	
10	0.772	
15	0.760	
20	0.740	
30	0.686	
50	0.509	
90	0.000	
Fertility rate per woman ($R(a)$), by age		[5]
0–14	0.000	
15–29	0.109	
30–49	0.300	
50+	0.000	
Population trimming	10% if population size exceeds 440.	Assumption
Transmission of infection		
General transmission parameters		
Relative biting rate (*rbr*)	$rbr = 1$; applies only to filarial transmission; not identified.	
Overall exposure rate of human hosts to central reservoir of infection (ζ)	Estimated from data (see main manuscript and Additional File 2); not identified, but used as main parameter to set level of transmission.	[6]
Seasonal variation in contribution to reservoir (*mbr*)	Stable throughout the year.	Assumption
Transmission probability (*v*)	$v = 1$; not identified.	
Success ratio (*sr*)	$sr = 1$; not identified.	
Zoophily (*z*)	$z = 0$; not applicable.	
Individual relative exposure to cloud		
Variation in by age and sex (*Exa*)	Linearly increasing from 0 to 1 between ages 0–10 and stable thereafter; no difference between males and females.	Assumption

Parameter	Value	Source
Variation due to personal factors (fixed through life) given age and sex (α_{Exi})	Estimated from data (see main manuscript and Additional File 2).	[6]
Individual relative contribution to cloud		
Variation by age and sex (*Coa*)	Linearly increasing from 0 to 1 between ages 0–10 and stable thereafter; no difference between males and females.	Assumption
Variation due to personal factors (fixed through life) given age and sex (α_{Coi})	Assumed to be perfectly correlated with individual exposure to reservoir, given age and sex (i.e. $Coi = Exi$).	Assumption
Host immunity to incoming infections		
Average impact of host immunity (α_{Imm})	$\alpha_{Imm} = 0$; i.e. no effect of immunity on incoming infections.	Assumption
Immunological memory (β_{Imm})	Irrelevant given that $\alpha_{Imm} = 0$.	Assumption
Life history and productivity of the parasite in the human host		
Average worm lifespan (*Tl*)	3 years	[7–9]
Variation in worm lifespan	Weibull distribution with shape 2; i.e. the mortality rate is zero at age zero and then increases linearly with age.	Assumption
Prepatent period (*pp*)	7 weeks	[10,7,8,11]
Age-dependent reproductive capacity (*R(a)*)	*R(a)* = 1 for patent female worms of any age.	Assumption
Longevity of infective material within host (*Tm*)	1 month; i.e. the minimum given that transmission is simulated in discrete time steps of one month.	Assumption
Mating cycle (*rc*)	1 month; i.e. the minimum given that transmission is simulated in discrete time steps of one month.	Assumption
Male potential (*pot*)	100 female worms.	Assumption
Density-dependent female worm reproductive capacity		
Worm contribution to host load of infective material in absence of density dependence (a_o)	$a_o = 200$ epg/worm; constant, i.e. no variation between human hosts.	[12]
Hyperbolic saturation: maximum total output of female worm population in a host (b_o)	Several alternative assumptions (see main text).	
Exponential saturation of individual female worm productivity per worm present in host (λ_z)	$\lambda_z = 0$, i.e. no exponential saturation; this type of saturation is too strong and causes a decline in total egg output once a certain number of worms is present in the host, such that we cannot reproduce distributions of infection intensities as observed in the field.	[13–19]
Morbidity		
Disease threshold (*Elc*)	Not used.	

Parameter	Value	Source
Reduction in remaining life expectancy due to disease (rl)	Not used.	
Infection dynamics in the cloud		
Cloud uptake of infectious material ($U(.)$)	The identity function, meaning there is no density-dependence in uptake of infective material by the environmental reservoir.	Assumption
Monthly cumulative survival of infective material in the central reservoir (ψ)	11.5%; $e^{\left(-\frac{52}{2}\frac{1}{12}\right)}$, assuming that survival of infective material is exponential and the average lifespan is two weeks (95%-CI: 0.05–7.38 weeks), based on the notion that average survival is a matter of "weeks" according to literature.	[10,11,20]
Mass treatment coverage		
Coverage (C_w)	User-defined.	
Relative compliance ($c_r(k,s)$) by age and sex (descriptive label used in graphs)		[21,22]
0–2, both sexes (infants)	0 (excluded from preventive chemotherapy)	
2–5, both sexes (preSAC)	1	
5–14, both sexes (SAC)	1	
15–44, males (other)	0 or 1 (only when targeting whole population)	
15–44, females (WCBA)	0 or 1 (only when targeting WCBA)	
45–90, both sexes (other)	0 or 1 (only when targeting whole population)	
Individual compliance index (co_i)	Uniform distribution [0,1].	[23]
Drug treatment		
Proportion of larvae or eggs cleared from host	0%	Assumption
Duration of temporary reduction in female reproductive capacity (Tr_0)	0 months	Assumption
Permanent reduction in female worm reproductive capacity (d_0)	0%	Assumption
Proportion of adult worms killed (m_0)	$m_0 = 0.95$ for albendazole and $m_0 = 0.80$ for albendazole.	[19]
Relative effectiveness (v)	$v = 1$ (constant, i.e. no additional variation).	Assumption
Vector control		
Timing	Not used.	
Coverage	Not used.	
Surveys		
Dispersal factor for worm contribution to measured density of infective material (d)	d = 1; constant, i.e. no additional variation; not identified.	

Parameter	Value	Source
Variability in measured host load of infective material (eggs per gram faeces)	Negative binomial distribution with mean $ss(t)$ and aggregation $k = 0.4$ (estimated from data, see Additional File 2); not identified, but used as main parameter to quantify variation in observed egg counts.	[24]

3. Formal description of WORMSIM (v2.58Ap9)

Human demography

The human population dynamics is governed by birth and death processes. We define $F(a)$ as the probability to survive to age a (Table A1-3). The cumulative survival for intermediate ages is obtained by linear interpolation.

The expected number of births (per year) at a given moment t is given by:

$$R_b(t) = \sum_{a=1}^{n_a} N_f(a,t) \cdot r_b(a) \tag{1}$$

with:

$N_f(a,t)$ number of women in age group a at time t

$r_b(a)$ annual birth rate in age-group a (Table A1-3).

n_a number of age-groups considered.

Each month, $R_b(t)$ is adapted according to the number of women and their age-distribution.

Once every year, the total number of human individuals is checked; if the total number is larger than a user-defined value, a fraction (also user-defined) is randomly removed from the simulation.

The population distribution resulting from the aforementioned parameters is illustrated in Figure A1-1, and closely follows the age distribution in Sub-Saharan Africa as estimated by the UN Population Division for the year 2000 (Figure A1-1) [5].

Figure A1-1. Population demography simulated in WORMSIM in absence of excess mortality due to disease (bars), compared to the 2000 population for Sub-Saharan Africa (diamonds; UN Population Division, World Population Prospects: The 2012 Revision).
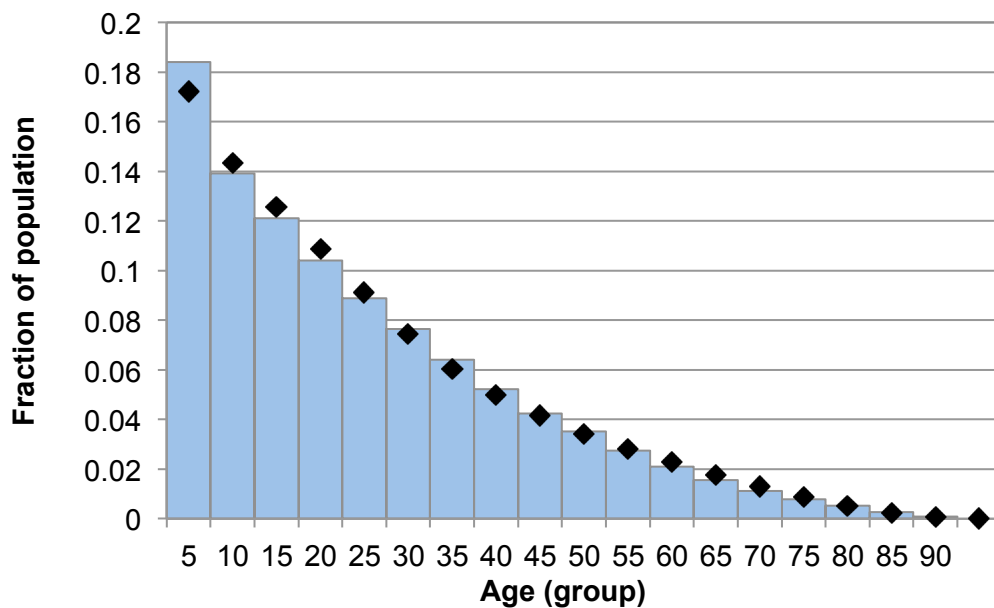
Transmission of infection

In WORMSIM, transmission between individuals is mediated by a conceptual cloud, which either represents a vector population or an environmental reservoir of infection. Individual human hosts are *exposed* to the infective material in the cloud at varying rates, given their age, sex, and personal factors. Vice versa, individual hosts *contribute* infective material (larvae or eggs) to the cloud, the amount depending on the number and reproductive statues of worms in the individual, as well as an individual host's contribution rate (depending on age, sex, and personal factors). The amount of infective material in the cloud is updated in discrete monthly time steps. Below, we describe how WORMSIM simulates a full transmission cycle: human exposure to infection and acquisition of worms, dynamics of infection within humans, contribution of infective material to the cloud, and within-cloud dynamics of infective material.

Exposure to infection and acquisition of new worms

First, we define the overall force of infection $lr(t)$ acting on the human population in month t as a function of the current absolute amount of infective material in the cloud $\overline{lu}(t)$:

$$lr(t) = \overline{lu}(t) \cdot \zeta \cdot v \tag{2}$$

Here, ζ (zeta) is a scalar representing the overall exposure rate, and v is the probability that an infective particle in the reservoir successfully develops into a parasite life stage that is capable of infecting a human host.[a]

Next, we define the force of infection acting upon individual i of age a and sex s as:

$$foi_i(t) = lr(t) \cdot \frac{Ex_i}{\sum_{i=1}^{N(t)} Ex_i} \tag{3}$$

Here, Ex_i is the relative exposure of an individual, taking into account age a and sex s, as well as personal factors:

$$Ex_i = Exa(a_i, s_i) \cdot Exi_i \tag{4}$$

with:

$Exa(a_i, s_i)$	Relative exposure of person with age a and sex s, defined as a linearly interpolated function of user-defined exposure rates for a finite set of ages (for each sex).
Exi_i	Exposure index of person i, which captures personal factors related to e.g. behaviour and occupation. Exi_i is assumed to follow a gamma distribution with mean 1.0 and shape and rate (or 1/scale) equal to α_{Exi}. The exposure index of a person remains constant throughout lifetime.[b]

---

[a] ζ is perfectly negatively correlated with transmission probability v, success ratio sr, relative biting rate rbr, and vector zoophily z. See also the section on contribution of infective material to reservoir. For filarial transmission, we set $\zeta = 1$, quantify v based on vector biology, set sr to a constant value, and calibrate transmission with rbr. For STH, we set $v = rbr = sr = 1$, and calibrate transmission with ζ, which has a more natural explanation in the STH context (exposure to the reservoir) than rbr.

[b] If desired, other continuous probability function can be chosen.

Finally, a person i is assumed to become infected in month m, according to a Poisson process with rate equal to $foi_i(t) \cdot sr \cdot Imm_i(t)$. Here, success ratio sr is a constant representing the probability that an inoculated infective particle will develop into a macroparasite.[a] Finally, $Imm_i(t)$ represents the impact of the host's immune response in montht t on incoming infections [25]:

$$Imm_i(t) = 1 - \alpha_{imm} \cdot imm_i \cdot W_{cum,i}(t) \tag{5}$$

$$W_{cum,i}(t) = W_i(t) + \beta_{Imm} \cdot W_{cum,i}(t-1)$$

Here, α_{imm} is the effect of immunity and imm_i is an individual host's capacity to elicit an immune response, drawn from a positive bounded probability distribution with mean one (e.g. a gamma distribution with equal shape and rate (1/scale) parameters). $W_{cum,i}(t)$ is the cumulatively experienced worm burden of host i in month t, $W_i(t)$ is the worm burden of host i in month t, and β_{Imm} represents the immunological memory span ($\beta_{Imm} = e^{-\ln(2)/\lambda_{imm}}$, where λ_{imm} is the desired half-life (in months) of the immunological response; vice vera $\lambda_{imm} = -\ln(2)/\ln(\beta_{Imm})$).

Within-host dynamics of infection

For convenience, in this section we drop the subscript i for individual humans. The lifespan of male and female parasites within human hosts is a random variable: $Tl \sim \text{Weibull}(\mu_{Tl}, \alpha_{Tl})$, with mean μ_{Tl} years and shape α_{Tl}.[c] Once parasites come of age (i.e. when they pass the prepatent age pp), female worms can start producing larvae or eggs, and males can inseminate female worms. The reproductive capacity $r(a,t)$ of a patent female worm of age a at time t is calculated as follows (in absence of drug effects):

$$r(a,t) = R(a - pp) \cdot m(t) \cdot z(t) \tag{6}$$

with:

$R(A)$ Potential reproductive capacity of a female worm, A years after reaching patency, defined as a linearly interpolated function of user-defined values for a finite set of ages.

$m(t)$ Mating factor at time t

$z(t)$ The exponential fecundity coefficient at time t, defined as $z(t) = e^{-W(t)\lambda_z}$, where $W(t)$ is the number of adult worms (males and females) in a given host at time t, and $\lambda_z \in \mathbb{R}^+$ quantifies the amount of negative density dependence. If $\lambda = 0$, there is no exponential saturation in egg production ($z(t) = 1$).

To produce larvae or eggs, a female worm must be inseminated each reproductive cycle rc, defined in terms of months. If insemination took place less than rc months ago, then $m(t) = 1$. Otherwise, the probability of insemination or reinsemination $P_{ins}(t)$ in month t is given by:

$$P_{ins}(t) = \begin{cases} pot \cdot W_m(t)/W_f(t) & \text{if } pot \cdot W_m(t) < W_f(t) \\ 1 & \text{if } pot < 0 \text{ or otherwise} \end{cases} \tag{7}$$

---

[c] For readers used to the other commonly used parameterization of the Weibull distribution in terms of shape k and scale λ, shape k is α_{Tl} (as described in this document) and scale $\lambda = \mu_{Tl}/\Gamma(1 + 1/\alpha_{Tl})$.

with:

$W(t)$ the number of male (W_m) or female (W_f) parasite in the human at time t

pot the number of female worms that a male worm can inseminate per month[d]

If no insemination takes place then $m(t) = 0$ and the female worm has a new opportunity to be inseminated in the next month $t + 1$. If insemination occurs in month t_i then $m(t) = 1$ during $t_i \le t < t_i + rc$.

The density of larvae (e.g. per skin snip) or eggs (e.g. per gram faeces) $sl(t)$ in a host at time t is calculated by accumulating the production of all female parasites over the past *Tm* months within that host:

$$sl(t) = O\big(el(t)\big) \tag{8}$$

$$el(t) = \sum_j^{n_i} d_j \sum_{x=1}^{Tm} r_j\big(a_j - x, t - x\big) \tag{9}$$

with:

$el(t)$ the *effective parasite load* at time t. This intermediate variable describes the female parasite load obtained by weighting each worm according to the mf-productivity during the past *Tm* months.

$O(.)$ A function that returns the total amount of infective material produced by female parasites. For soil-transmitted helminths, we assume that $O(.)$ is the hyperbolic saturating function $cax/(1 + ax/b)$, where x is the number of worms, and a and b are shape parameters, and c is a scale parameter.[e]

d_j *dispersal factor* of female parasite j. This is a random variable (mean 1.0) drawn for every "newborn" worm, and accounts for differences in the contribution of female worms to the density at the standard site of the body where samples are taken or vectors bite.

Tm (fixed) lifespan of larvae or eggs within the host in terms of months.

n_i number of parasites alive during at least one of the months *t-1,…,t-Tm*.

Host contribution of infective material to the cloud

Given the density of larvae or eggs $sl_i(t)$ in all $N(t)$ host in month t, the total amount of infective material that is contributed to the cloud by the host is defined as

$$\overline{lu}(t)_{in} = \sum_{i=1}^{N(t)} Mbr(t) \cdot rbr \cdot U(sl_i(t)) \cdot Co_i \tag{10}$$

---

[d] When the user specifies a negative value for *male potential*, female worms can produce larvae or eggs in the absence of male worms.

[e] Alternatively, a linear or other functional relationship between *el* and *sl* can be defined. Saturating functions should not be used when $Tm > 1$, as this will cause partial saturation of female worm productivity in month t, given the output in months $t - 1$ through $t - Tm$. This will be alleviated in a future version of WORMSIM by setting $sl(t) = \sum_{x=1}^{Tm} O\big(\sum_j^{n_i} r_j(a_j - x, t - x)\big)$.

Here, $Mbr(t)$ is the average contribution rate in month t (*monthly biting rate* for filarial infections), allowing the user to define a seasonal pattern (in absence of vector control). The relative biting rate *rbr* is used to scale this seasonal pattern to some desired level.[a] The function $U(.)$ returns the amount of infective material taken up by the cloud given the density of eggs or larvae $sl_i(t)$ in a host, possibly in a density dependent manner to represent e.g. limited vectorial capacity to transmit infection.[f] Last, Co_i is the relative contribution of an individual, given age, sex, and personal factors:

$$Co_i = Coa(a_i, s_i) \cdot Coi_i \tag{11}$$

with:

$Coa(a_i, s_i)$	Relative contribution of person with age a and sex s, defined as a linearly interpolated function of user-defined exposure rates for a finite set of ages (for each sex).
Coi_i	Contribution index of person i, which captures personal factors related to e.g. behaviour and occupation. Coi_i is assumed to follow a gamma distribution with mean 1.0 and shape and rate (or 1/scale) equal to α_{Coi}. The contribute index of a person remains constant throughout lifetime. In WORMSIM default assumption is that $Coi_i = Exi_i$, unless separate distributions are defined.

Dynamics of infective material in the cloud

For the dynamics of infective material in the cloud we define a deterministic, discrete model:

$$\overline{lu}(t) = \overline{lu}(t)_{in} + \psi \cdot \overline{lu}(t-1) \tag{12}$$

Each month, new infective material $\overline{lu}(t)_{in}$ is added to the cloud, and a fixed proportion ψ of the infective material from the previous month is carried over, assuming exponential survival of infective material. The average life span of infective material in the cloud is then defined as $-1/\ln(\psi)$ months. To simulate filarial transmission, we set $\psi = 0$, such that the cloud represents a vector population in which larvae survive for much shorter than a month. To simulate hookworm or other STH infections, we set $0 < \psi < 1$, such that the cloud represents an environmental reservoir of infection in which infective material survives for a non-negligible time.

Morbidity

The event of a person developing symptoms at age a depends on the *accumulated parasite load (elc)* of a person:

$$elc(a) = \sum_{x=0}^{a} el(x) \tag{13}$$

Each person has a threshold level elc (denoted as Elc) at which a person goes blind. Elc follows a probability distribution: $Elc \sim \text{Weibull}(\mu_{Elc}, \alpha_{Elc})$, with mean μ_{Elc} and shape α_{Elc}. Person i goes blind at age a when:

$$elc_i(a) \geq Elc_i > elc_i(a-1) \tag{14}$$

---

[f] For filarial infection, $U(.)$ typically is a density-dependent function of $sl_i(t)$ to represent limited vectorial capacity to transmit infection, whereas for STH, we take $U(.)$ to be the identity function.

At that moment the remaining lifespan at age a is reduced by a factor rl which follows a user-defined distribution on $[0,1]$.[g]

Mass treatment coverage and compliance

The primary characteristic of a certain ivermectin mass treatment w is the coverage C_w (fraction of the population treated). However, a difficulty in calculating individual chances of participation is that there are several exclusion criteria for the drug. Moreover, compliance to treatment differs from person to person. Exclusion criteria can be either permanent (chronic illness) or transient (e.g. related to age or pregnancy). We define the population that potentially participates as the total population *minus* a fraction f_c that never participates in mass drug administration. The coverage among the potentially participating population C'_w is now given by:

$$C'_w = C_w/(1 - f_c) \tag{15}$$

Here, C'_w cannot be larger be than one (i.e. is capped off at one).

To capture transient contra-indications and other age- and sex-related factors for participation in mass treatment, we define the age- and sex-specific relative compliance $c_r(k,s)$ (Table A1-3). Note that in $c_r(k,s)$ only the *ratio* between the values for the different groups is relevant.

Now, the coverage $c(k,s,w)$ in each of the age- and sex-groups (among people that potentially participate) at treatment round w is calculated as:

$$c(k,s,w) = \frac{c_r(k,s) \cdot N(w)}{\sum_{s=1}^{2} \sum_{k=1}^{n_a} c_r(k,s) \cdot N(k,s,w)} \cdot C'_w \tag{16}$$

with:

$N(k,s,w)$

> Number of individuals eligible to treatment in age-group k and sex s at treatment round w.

$N(w)$ Total number of eligible individuals at treatment round w.

Finally, the probability to participate in treatment round w for an person i of age-group k and sex s is given by:

$$Ptr_{i,w} = co_i^{\frac{1-c(k,s,w)}{c(k,s,w)}} \tag{17}$$

with:

co_i Personal compliance index. This is considered as a lifelong property and is generated by a uniform distribution on $[0,1]$

Note that for all k and s the average value of $Ptr_{i,w}$ equals $c(k,s,w)$. Now, in WORMSIM we define 3 coverage models. In model 0, the probability to be treated is as given in equation (17). In model 1, the probability is equal to $c(k,s,w)$ and the compliance index co_i is ignored. The simplest model is model 2 in which the treatment probability simply equals C'_w. All models take account of a fraction f_c of permanently excluded persons. Figure A1-2

---

[g] For STH modelling, we do not use this feature, and thus assume zero excess mortality from infection.

illustrates the impact of different assumptions about compliance patterns on the proportion of the population that has never been treated after a certain number of treatment rounds.

Figure A1-2. Relation between compliance patterns and proportion of population that has never been treated. For simplicity, here we assume that compliance is independent of age and sex. Random compliance (solid line) means that eligible individuals participate completely at random (compliance model 1 or 2 in WORMSIM, depending on whether age and sex-patterns are required). Systematic compliance (dotted line) means that an individual either always participates (if eligible) or never (compliance model 1 or 2 in WORMSIM, combined with a fraction of excluded people equal to one minus the target coverage). The mixed compliance pattern (dashed line) means that some individuals are systematically more likely to participate than others (but everyone will participate at some point; compliance model 0 in WORMSIM).
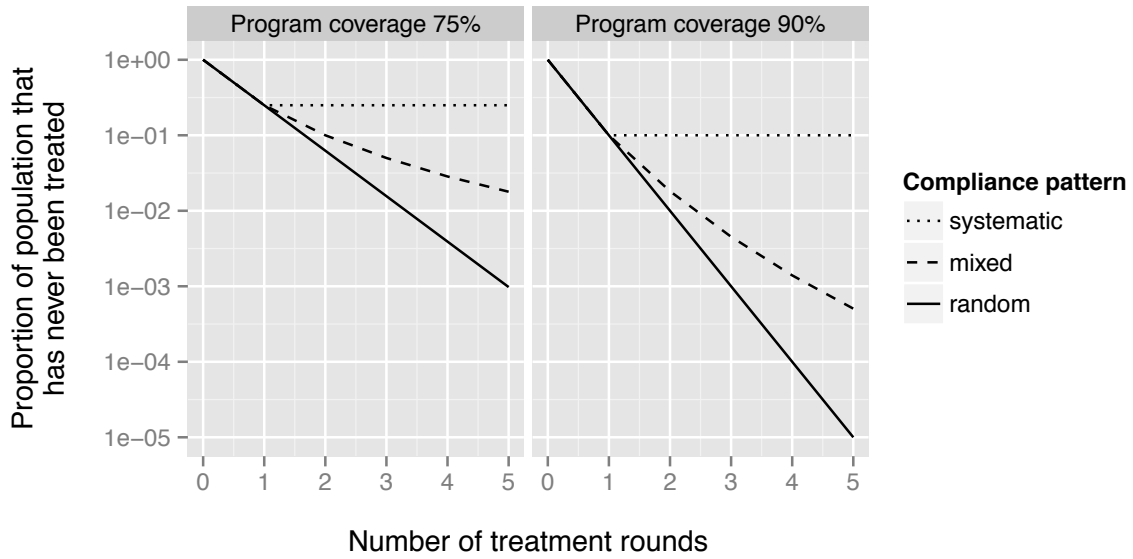
Parasitological effects of treatment

In WORMSIM, drug treatment affects parasites in three main ways. First, a drug may instantly kill a proportion of larvae or eggs present in a host. This proportion is either fixed or a randomly drawn from a user-defined probability distribution for each host and treatment.

Second, a drug may instantly kill pre-patent and adult worms with probability m_i in host i. A worm j dies when a random variate u_j on [0,1] (redrawn for every new treatment) is smaller than or equal to m_i.

Third, a drug may temporarily and/or permanently (and cumulatively) reduce the reproductive capacity of female worm by a proportion d_i in host i. In case of a temporary effect, the reproductive capacity will restore within a period Tr_i to its maximum value (in case of any concomitant permanent reductions, reproductive capacity will regenerate to the new, permanently reduced maximum value. The second and third effect are jointly defined as follows:

$$m_i = v_i m_0$$
$$d_i = v_i d_0 \hspace{4cm} (18)$$
$$Tr_i = v_i Tr_0$$

$$r_j(a_j, t) = r_j^0(a_j, t) \cdot (1 - d_i) \cdot \left(\frac{\tau}{Tr_i}\right)^s, \quad \text{if } u_j > m_i, d_i < 1, \text{ and } t < Tr_i$$

$$r_j(a_j, t) = r_j^0(a_j, t) \cdot (1 - d_i) \qquad \text{if } u_j > m_i, d_i < 1, \text{ and } t \geq Tr_i$$

$$r_j(t) = 0 \qquad\qquad\qquad\qquad \text{otherwise}$$

with:

v_i Relative effectiveness of treatment in person i. For every separate treatment and person, a new value is drawn for v_i from a user-defined distribution (i.e. the relative effectiveness applies to all worms in a person during a specific treatment).

m_0 Average fraction of prepatent and adult parasites killed.

d_0 Average permanent (unrecoverable) reduction in female reproductive capacity.

Tr_0 Average duration until full recovery from temporary effects on female reproductive capacity.

$r_j(a_j, t)$

     Reproductive capacity of female worm j in month t, τ months after the last treatment.

$r_j^0(a_j, t)$

     Reproductive capacity of female worm j had person i not been treated at the last round, τ months ago.

s Shape parameter of the recovery function.

In addition to this, we explicitly consider that some persons (a user-defined random fraction of the treated population) do not at all react to the drug during a certain treatment due to malabsorption (e.g. due to vomiting or diarrhoea).

Vector control

Vector control is modelled as a reduction of the monthly biting rates during a given period of time. A period of vector control[h] is specified as the year + month of the beginning of the strategy and the year + month of the end of a strategy. If a certain month during a period of d days larvicides have been applied, then the reduction in $Mbr(t)$ in that month equals $d/30$ x 100%.

Surveys

During the simulation, user-defined surveys will take place. During a survey, for all simulated individuals the actual number of male and female worms is recorded, and a diagnostic test is simulated to detect infective material (larvae, eggs). For the diagnostic test, the expected amount of infective material per sample (e.g. microfilariae per skin snip, or eggs per gram faeces) for an individual is given by $sl_i(t)$.

The actual number of infective particles (microfilariae, eggs, etc.) in the sample is assumed to follow a discrete distribution like a Poisson or negative binomial distribution, with mean

---

[h] Multiple periods of vector control can be specified, each with its own effectiveness.

equal to $sl_i(t)$.[i] At each epidemiological survey a user-defined number of samples are taken from all simulated persons, for which the results are averaged (per simulated person). The results of such a survey are post-processed to arrive at age and sex-specific prevalences and intensities of infection.

Simulation warm-up

In general, before starting simulation of interventions in ONCHOSIM, a 200-year warm-up period is simulated, such as to allow the human and worm population to establish equilibrium levels, given the parameters for average fly biting rate and inter-individual variation in exposure to infection. At the start of the warm-up period, an artificial force of infection is simulated for a user-defined number of years, allowing worms to establish themselves in the human population (here: 5 worms per person per year for 5 years). After the 200 warm-up years, the simulated infection levels are no longer correlated with the initial conditions at the start of the warm-up period.

---

[i] For filariasis transmission, we typically assume that sampling error is Poisson distributed. For faecal egg counts in STH infection, we assume that sampling error is negative binomial, while setting d_j=1. For hookworm, we assume that aggregation parameter $k = 0.4$, based on an analysis of field data [24], kindly provided by Simon Brooker.

4. Instructions for installing and running WORMSIM

Installing WORMSIM

Download and install the Java SE Runtime Environment 8 from
http://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html

Download and unzip wormsim-2.58Ap9.zip to a location of your choice on your computer (Additional File 2).

A folder named wormsim-2.58Ap9 will be created that contains:

- a number of example XML input files (all ending in .xml)
- *wormsim.xsd*, the XML Schema that is used to validate input files
- *wormsim.jar*, a Java archive with the .class and .java files of WORMSIM
- *colt.jar*, the Colt library by Wolfgang Hoschek (CERN) that is used for statistical distributions
- *run.sh* and *run.bat*, a script / batch file to run WORMSIM
- *avg.sh* and *avg.bat*, a script / batch file to aggregate the output of individual runs produced by running WORMSIM
- *test.sh*, an example script / batch file that calls run.sh/run.bat and avg.sh/avg.bat
- *readme.txt*, a text file documenting the history of changes to WORMSIM
- *license.txt*, a test file describing the license and conditions for using WORMSIM

Running WORMSIM

Microsoft Windows

Test the successful installation by:

- opening a DOS command line window by clicking on Start (Windows 7) and typing `cmd` and pressing enter. Navigate to the folder where WORMSIM has been installed, for instance (if you downloaded the zip file to your desktop and unzipped in that location):
  ```
  cd .\Desktop\wormsim-2.58Ap9
  ```
- if you do not have any experience with running batch files, you will find a tutorial at
  http://www.computerhope.com/issues/chusedos.htm
- running the test.bat batch file by typing:
  ```
  .\test
  ```
- after running *test.bat* you should find the following files in your WORMSIM folder:

  example_STH.log

  example_STH.txt

  example_STH0-19.zip, a zip file containing output of individual runs

- **see the supplement WORMSIM output documentation for details about the output files**

Copy the *example_STH.xml* input file and edit this file for your specific scenario. To run WORMSIM with the new input file, copy and edit the *test.bat* batch file. Assuming you

copied *example_STH.xml* to *my_STH.xml* and *test.bat* to *my_test.sh*, you would edit the contents of the new *my_test.sh* as follows:

```
.\run.bat my_STH.xml 0 99
.\avg.bat my_STH.xml 0 99
```

and run your shell script with:

```
.\my_test
```

to do 100 runs and aggregate the output of these runs.

Mac OS X or Linux

Test the successful installation by:

- opening a Terminal window by running the Terminal program (to be found in Utilities) and navigating to the folder where Wormsim has been installed, for instance (if you downloaded the zip file to your desktop and unzipped in that location):
  ```
  cd ~/Desktop/wormsim-2.58Ap9
  ```
- if you do not have any experience with running shell scripts, you will find an excellent Unix/Linux tutorial at http://www.ee.surrey.ac.uk/Teaching/Unix/
- running the test.sh shell script by typing:
  ```
  ./test.sh
  ```
- after running test.sh you should find the following files in your Wormsim folder:

  example_STH.log

  example_STH.txt

  example_STH0-19.zip, a zip file containing output of individual runs
- see the supplement Wormsim output documentation for details about the output files

Copy the *example_STH.xml* input file and edit this file for your specific scenario. To run WORMSIM with the new input file, copy and edit the *test.sh* shell script. Assuming you copied *example_STH.xml* to *my_STH.xml* and *test.sh* to *my_test.sh*, you would edit the contents of the new *my_test.sh* as follows:

```
./run.sh my_STH.xml 0 99
./avg.sh my_STH.xml 0 99
```

and run your shell script with:

```
./my_test.sh
```

to do 100 runs and aggregate the output of these runs

Output options

The -d output option will make WORMSIM produce additional detailed output. This output is found in **X.txt* and **Y.txt* (for instance *example_STHX.txt* and *example_STHY.txt*).

The −n output option suppresses all output except the **.log* output (e.g. *example_STH.log*).
Either output option can be added to the *run* command as follows:

```
./run.sh my_STH.xml 0 99 -d
```

or
```
./run.sh my_STH.xml 0 99 -n
```

5. Annotated input file

The WORMSIM inputfile is an XML file that can be edited with any text editor or alternatively, with an XML editor (such as Oxygen XML Editor).
The advantage of using the XML format is that any input file can be validated against an XML Schema (a formal specification of the grammar used in the specific XML dialect used for the WORMSIM input file).

The Wormsim input file is document with an annotated example (annotated-example.xml) and by the overview below. The XML Schema (wormsim.xsd) is documented in great detail in schema-documentation-wormsim.pdf (provided within Additional File 4).

The main elements of the Wormsim inputfile are:
- `simulation`
- `demography`
- `blindness`
- `exposure.and.contribution`
- `immunity`
- `worm`
- `fly`
- `mass.treatment`
- `vector.control`

We will cover each of these elements in more detail below.

<simulation>

The <simulation> element specifies the start year of the simulation, the timing of surveys (i.e. output moments), the number of skin snips taken at each survey and the age classes for output.

The comments (text formatted as `<!— this is a comment -->`) in the input file fragment below give more detail.

```
<simulation start.year="1800">
 <!-- number of skin snips taken per person                 -->
 <surveillance nr.skin-snips="1" skin-snip.categories="0,1999,3999,1e6>

 <!-- timing of surveys                              -->
 <!-- month 0 represents January 1st                 -->
 <!-- see note regarding "delay" below               -->
  <periodic.surveys>
   <start year="2000" month="0" delay="-2"/>
   <stop year="2015" month="1"/>
   <interval years="1" months="0"/>
  </periodic.surveys>
  <extra.surveys>
   <survey year="2011" month="0" delay="-2"/>
   <survey year="2012" month="0" delay="-2"/>
   <survey year="1810" month="0" delay="-2"/>
   <survey year="1820" month="0" delay="-2"/>
   <survey year="1830" month="0" delay="-2"/>
   <survey year="1850" month="0" delay="-2"/>
   <survey year="1900" month="0" delay="-2"/>
  </extra.surveys>
  <!-- upper bounds of age categories in output      -->
  <age.classes>
   <age.class age.limit="2"/>
   <age.class age.limit="5"/>
   <age.class age.limit="10"/>
   <age.class age.limit="15"/>
   <age.class age.limit="20"/>
   <age.class age.limit="30"/>
   <age.class age.limit="45"/>
   <age.class age.limit="90"/>
  </age.classes>
 </surveillance>
 <!-- upper bounds and weights of reference population for -->
 <!-- age and sex-standardized output (OCP standard pop)   -->
 <!-- in literature, the weight for age <5 is usually zero -->
 <standard.population>
  <age.group age.limit="5" n.males="1401" n.females="1353"/>
  <age.group age.limit="15" n.males="3508" n.females="2972"/>
  <age.group age.limit="25" n.males="1790" n.females="1790"/>
  <age.group age.limit="50" n.males="3092" n.females="3690"/>
  <age.group age.limit="90" n.males="1208" n.females="1237"/>
 </standard.population>
</simulation>
```

<demography>

The <demography> element defines life tables for the male and female population, the maximum population size (above which random persons will be removed), a fertility table, and the initial population size and age distribution. See comments below.

```
<!-- demographic parameters of simulated population              -->
<demography>
 <!-- whenever the simulated population size exceeds the    -->
 <!-- the specified maximum, a random fraction is removed   -->
 <!-- see note regarding "delay" below                      -->
 <the.reaper max.population.size="440" reap="0.1" delay="-3"/>
 <!-- survival represents cumulative survival probability   -->
 <!-- and is determined by for unspecified ages by linear   -->
 <!-- interpolation of values for specified age limits       -->
 <life.table>
  <survival age.limit="5" male.survival="0.804" female.survival="0.804"/>
  <survival age.limit="10" male.survival="0.772" female.survival="0.772"/>
  <survival age.limit="15" male.survival="0.760" female.survival="0.760"/>
  <survival age.limit="20" male.survival="0.740" female.survival="0.740"/>
  <survival age.limit="30" male.survival="0.686" female.survival="0.686"/>
  <survival age.limit="50" male.survival="0.509" female.survival="0.509"/>
  <survival age.limit="90" male.survival="0.000" female.survival="0.000"/>
 </life.table>
 <!-- rates represent probabilities for women to give birth  -->
 <!-- to one child in some year, given a woman's age         -->
 <!-- rates are assumed constant within each age category   -->
 <!-- and ages limits represent upper bounds of categories -->
 <!-- see note regarding "delay" below                      -->
 <fertility.table delay="-4">
  <fertility age.limit="5" birth.rate="0"/>
  <fertility age.limit="10" birth.rate="0"/>
  <fertility age.limit="15" birth.rate="0"/>
  <fertility age.limit="20" birth.rate="0.109"/>
  <fertility age.limit="30" birth.rate="0.300"/>
  <fertility age.limit="50" birth.rate="0.119"/>
  <fertility age.limit="90" birth.rate="0.0"/>
 </fertility.table>
 <!-- population size to start simulation with              -->
 <initial.population>
  <age.group age.limit="1" n.males="9" n.females="10"/>
  <age.group age.limit="5" n.males="29" n.females="34"/>
  <age.group age.limit="10" n.males="33" n.females="32"/>
  <age.group age.limit="15" n.males="25" n.females="24"/>
  <age.group age.limit="20" n.males="21" n.females="21"/>
  <age.group age.limit="30" n.males="35" n.females="35"/>
  <age.group age.limit="50" n.males="44" n.females="32"/>
  <age.group age.limit="90" n.males="19" n.females="20"/>
 </initial.population>
</demography>
```

<blindness>

The <blindness> element defines the parameters for development of morbidity ("blindness" as originally developed for ONCHOSIM, where we specify a threshold of cumulative exposure to microfilaria) and the effect of morbidity on the remaining life expectancy. See comments below.

```
<!-- parameters for development of blindness              -->
<!-- when a person's cumulative exposure to mf exceeds a thres- -->
<!-- hold, a person is considered blind                   -->
<!-- individual variation in susceptibility is modeled by let-  -->
<!-- ting the threshold vary between individuals, assuming a    -->
<!-- continuous distribution with some mean and shape "p1",     -->
<!-- optionally truncated at specified bounds "min" and "max"  -->
<blindness>
 <treshold dist.nr="0" mean="1"/>
 <!-- upon developing symptoms, the life-expectancy of a  -->
 <!-- person is reduced by a variable fraction, which is  -->
 <!-- drawn from a distribution on the domain [0,1]       -->
 <pct-life-expectancy-reduction dist.nr="0" mean="0"/>
</blindness>
```

\<exposure.and.contribution\>

The `<exposure.and.contribution>` element defines the parameters for the exposure of humans to a vector (or infectious reservoir) and the contribution of humans to the vector cloud (or infectious reservoir). See comments below.

```xml
<exposure.and.contribution>
 <!-- parameters for exposure to fly bites              -->
 <!-- N.B. in WORMSIM we only describe fly bites on humans  -->
 <environment zeta="0.528" psi="0.1145588"/>
 <!-- initial force of infection to introduce infection   -->
 <!-- into the simulated population; duration in years     -->
 <initial.foi duration="5" foi="5"/>
 <!-- parameters for individual exposure to fly bites,    -->
 <!-- depending on gender, age, and personal factors      -->
 <male>
  <!-- age-dependent exposure, relative to exposure -->
  <!-- of individuals with relative exposure one,   -->

  <!-- assuming a linear increase between age 0 and -->
  <!-- 10, after which exposure is 1.0              -->
  <exposure.function fun.nr="1" a="0.1" c="1.0"/>
  <!-- individual variation in exposure related to  -->
  <!-- e.g. occupation and attractiveness to flies, -->
  <!-- assuming a gamma distribution with mean one  -->
  <!-- and variation 1/p1 (shape and rate p1),      -->
  <!-- truncated by "min" and "max"                 -->
  <exposure.index dist.nr="4" min="0" max="20" p1="0.844"/>
  <!-- age-dependent exposure, relative to mean exp -->
  <!-- of adult males, assuming a linear increase   -->
  <!-- between age 0 and 10, after which exposure is 1.0 -->

  <contribution.function fun.nr="1" a="0.05" c="1"/>
  <!-- individual variation in contribution related to  -->
  <!-- e.g. occupation and attractiveness to flies, -->
  <!-- assuming a gamma distribution with mean one  -->
  <!-- and variation 1/p1 (shape and rate p1),      -->
  <!-- truncated by "min" and "max"; if not speci-  -->

  <!-- fied, an individual's contribution index is  -->

  <!-- taken to be equal to his/hers exposure index -->

  <!-- (the line below is commented out for STH)    -->
  <!--contribution.index dist.nr="4" min="0" max="20" p1="0.844"/-->
 </male>
 <female>
  <exposure.function fun.nr="1" a="0.1" c="1.0"/>
  <exposure.index dist.nr="4" min="0" max="20" p1="0.844 "/>
  <contribution.function fun.nr="1" a="0.1" c="1.0"/>
  <!--contribution.index dist.nr="4" min="0" max="20" p1="0.844 "/-->
 </female>
</exposure.and.contribution>
```

\<immunity\>

The `<immunity>` element defines the (optional) development of host immunity.
See comments below.

```
<!-- parameters related to development of host immunity against -->
<!-- incoming infections; these are currently set such that no  -->
<!-- immunity develops                                          -->
<immunity>
 <male alpha="0" beta="1">
 <immunity.function fun.nr="0" a="1"/>
 <immunity.index dist.nr="0" min="0" max="20"/>
 </male>
 <female alpha="0" beta="1">
  <immunity.function fun.nr="0" a="1"/>
  <immunity.index dist.nr="0" min="0" max="20"/>
 </female>
</immunity>
```

\<worm\>

The \<worm\> element defines parameters for worm lifespan, prepatent period, mating between M and F worms, age-dependent production of microfilaria (or eggs), mf (egg) density per worm and skin dispersal. See comments below for details

```
<!-- parameters for worm survival and mf production      -->
<!-- mf lifespan in months, see note regarding "delay" below   -->
<worm mf-lifespan="1" monthly.event.delay="+1">
 <!-- worm lifespan in years, allowing for variation      -->
 <!-- between worms, assuming a Weibull distribution with -->
 <!-- mean 10 and shape 3.76, bounded by "min" and "max"  -->
 <lifespan dist.nr="3" mean="3" p1="2"/>
 <!-- pre-patent period during which worms do not produce -->
 <!-- mf                                                  -->
 <prepatent.period dist.nr="0" mean="0.1346"/>
 <!-- number of months a female can produce mf with one   -->
 <!-- insemination, and number of females one male worm   -->
 <!-- can inseminate per month                            -->
 <!-- if there are more female worms than the total male  -->
 <!-- potential, every female has a probability of being  -->
 <!-- inseminated equal to N_mw/N_fm*male.potential       -->
 <mating cycle="1" male.potential="100"/>
 <!-- mf production by female worms as function of worm    -->
 <!-- age minus pre-patent period; mf production at un-    -->
 <!-- specified ages is determined by linear interpol;    -->

 <!-- optionally, labda represents the exponential        -->

 <!-- decline in female worm fecundity as a function of   -->

 <!-- total number of worms (males and females) in a host -->
 <age.dependent.mf-production labda="0">
  <mf-production age.limit="0" production="1"/>
  <mf-production age.limit="10" production="1"/>
  <mf-production age.limit="20" production="0"/>
 </age.dependent.mf-production>
 <!-- expected N_mf (or eggs) per worm in sample per       -->

 <!-- fully fecund worm (onchocerciasis configuration)     -->

 <!--skin.mf-density.per.worm fun.nr="1" a="7.6" b="0 " c="-1"/-->

 <!-- expected N_mf (or eggs) per worm in sample per       -->

 <!-- fully fecund worm (STH configuration)                -->

 <alt.skin.mf-density.per.worm>
  <a dist.nr="0" mean="200"/>
  <b dist.nr="4" mean="1500" p1="50"/>
  <c dist.nr="0" mean="1"/>
 </alt.skin.mf-density.per.worm>

 <!-- random dispersal factor for mf per worm        -->
 <skin.dispersal dist.nr="0" p1="1"/>
 <!-- NB distribution for observed number of mf in        -->
 <!-- one Kato-Katz slide of 41.7 mg                      -->
 <skin-snip.variability dist.nr="4" p1="0.40"/>
</worm>
```

\<fly\>

The `<fly>` element defines parameters that determine the successful uptake and development of L1 larvae into infective L3 larvae and also determins the fly biting rate. See below for details.

```
<!-- probability that an mf or egg/larva in the       -->
<!-- reservoir develops into an infectious particle    -->
<fly transmission.probability="1.0">
 <!-- functional relation between uptake of mf/eggs and -->
 <!-- mf/egg density in the host                        -->
 <L1-uptake fun.nr="1" a="1.0" b="0.0" c="-1"/>
 <!-- seasonal pattern in transmission (optional)       -->
 <monthly.biting.rates relative.biting.rate="0.4">
  <mbr month="1"  rate="1"/>
  <mbr month="2"  rate="1"/>
  <mbr month="3"  rate="1"/>
  <mbr month="4"  rate="1"/>
  <mbr month="5"  rate="1"/>
  <mbr month="6"  rate="1"/>
  <mbr month="7"  rate="1"/>
  <mbr month="8"  rate="1"/>
  <mbr month="9"  rate="1"/>
  <mbr month="10" rate="1"/>
  <mbr month="11" rate="1"/>
  <mbr month="12" rate="1"/>
 </monthly.biting.rates>
</fly>
```

<mass.treatment>

The <mass.treatment> element defines parameters for the timing of mass treatment rounds, individual compliance (permanent, temporary and age dependent), and effects of ivermectin on mature worms, mf production by F worms and on mf.

```
<mass.treatment>
 <!-- timing of individual mass treatment rounds (one line per       -->
 <!-- mass treatment round), specifying year, month (0 represents    -->
 <!-- January 1st), and population coverage (fraction of total       -->
 <!-- village population, including those not eligible for treatment) -->
 <!-- see note regarding "delay" below                               -->
 <treatment.rounds>
  <treatment.round year="2016" month="0" coverage="0.498" delay="-1"/>
  <treatment.round year="2016" month="6" coverage="0.498" delay="-1"/>
  <treatment.round year="2017" month="0" coverage="0.498" delay="-1"/>
 </treatment.rounds>
 <!-- random fraction of population permanently not eligible for -->
 <!-- treatment due to chronic illness and random fraction of    -->
 <!-- population in which the drug does not work due to diarrhoe -->
 <!-- (temporary effect)                                         -->
 <compliance fraction.excluded="0.0" fraction.malabsorption="0.0">
  <!-- weights for age and sex-specific compliance, given        -->
  <!-- some expected overall coverage in the eligible            -->
  <!-- population weights are constant within age categories     -->
  <age.and.sex.specific.compliance age.limit="2" male.compliance="0"
   female.compliance="0"/>
  <age.and.sex.specific.compliance age.limit="5" male.compliance="1"
   female.compliance="1"/>
  <age.and.sex.specific.compliance age.limit="10" male.compliance="1"
   female.compliance="1"/>
  <age.and.sex.specific.compliance age.limit="15" male.compliance="1"
   female.compliance="1"/>
  <age.and.sex.specific.compliance age.limit="20" male.compliance="0"
   female.compliance="1"/>
  <age.and.sex.specific.compliance age.limit="30" male.compliance="0"
   female.compliance="1"/>
  <age.and.sex.specific.compliance age.limit="45" male.compliance="0"
   female.compliance="1"/>
  <age.and.sex.specific.compliance age.limit="90" male.compliance="0"
   female.compliance="0"/>
 </compliance>
 <!-- drug efficacy, specified as permanent reduction in worm     -->
 <!-- capacity to produce mf (cumulative effects allowed), pattern -->
 <!-- of how mf production recovers over time (to a new, reduced  -->
 <!-- maximum level), and fraction of mf surviving each treatment -->
 <treatment.effects permanent.reduction.mf-production="0.0 "
  period.of.recovery="0.01" shape.parameter.recovery.function="1.0"
  fraction.killed="0.95">
  <fraction.mf.surviving dist.nr="0" mean="1.0"/>
  <!-- variability in treatment effects (relative to mean expected -->
  <!-- effect)                                                     -->
  <treatment.effect.variability dist.nr="0" mean="1.0"/>
 </treatment.effects>
</mass.treatment>
```

\<vector.control\>

The `<vector.control>` element defines parameters for setting the effectivity of vector control during periods of vector control.

```
<!-- effectivity is specified as relative reduction in biting rates -->
<!-- if no vector control is desired in the simulation, set both    -->
<!-- start.year and stop.year outside the scope of the simulation   -->
<!-- (i.e. after the last survey)                                   -->
<!-- vector control is usually assumed to be highly effective       -->
<vector.control>
 <period start.year="2010" stop.year="2160" effectivity="0.95"/>
 <period start.year="2020" stop.year="2170" effectivity="0.95"/>
</vector.control>
```

6. Annotated output files

The WORMSIM **standard output** (e.g *example_STH.txt*) is a tab delimited text file with the following columns:

year	: time (years)
N	: nr examined
N20	: nr examined > 20 yrs old
mf+	: percentage with positive skin snip
mf+20	: percentage with positive skin snip (> 20 yrs)
mfPr	: age/sex standardized mf prevalence
aNmf	: arithmetic mean nr mf per skin snip
aNmf20	: arithmetic mean nr mf per skin snip (> 20 yrs)
gNmf	: geometric mean nr mf per skin snip
CMFL	: geometric mean nr mf per skin snip (> 20 yrs)
bl	: percentage blind
bl20	: percentage blind (> 20 yrs)
blPr	: age/sex standardized prevalence of blindness
w+	: percentage with at least one adult female worm
w+20	: percentage with at least one adult female worm (> 20 yrs)
aNw	: arithmetic mean nr of adult female worms per person
aNw20	: arithmetic mean nr of adult female worms per person (>20 yrs)
mbr	: monthly fly biting rate in previous month
mtp	: monthly transmission potential in previous month
L1	: mean nr of L1 larvae per 1000 biting flies in previous month
L3	: mean nr of L3 larvae per 1000 biting flies in previous month
foi	: mean force of infection (nr of new adult worms per person) in prev. month

The WORMSIM **log output** (e.g *example_STH.log*) is a tab delimited text file with output for each simulation run in the following columns:

seed	: the seed of the random number generator (i.e. run nr) of that specific run
year	: time (years)
mf+	: fraction with positive skin snip
mf5+	: fraction with positive skin snip > 5 yrs
w+	: fraction with at least one adult female worm
N	: nr examined
aNmf	: arithmetic mean nr mf per skin snip
aNmf20	: arithmetic mean nr mf per skin snip (> 20 yrs)
N20	: nr examined > 20 yrs old
CMFL	: geometric meain nr mf per skin snip (> 20 yrs)

The detailed *X.txt output (e.g *example_STHX.txt*) contains sex and age specific output with the following columns:

year   : time (years)
age   : upper limit of age group
M   : nr M examined in that age group
F   : nr F examined in that age group
Mbl   : percentage blind M in that age group
Fbl   : percentage blind F in that age group
Mmfpos   : percentage M in that age group with positive skin snip
Fmfpos   : percentage F in that age group with positive skin snip
MaNmf   : arithmetic mean nr mf per skin snip in M of that age group
FaNmf   : arithmetic mean nr mf per skin snip in F of that age group
MgNmf   : geometric mean nr mf per skin snip in M of that age group
FgNmf   : geometric mean nr mf per skin snip in F of that age group
Mwpos   : percentage M in that age with at least one adult female worm
Fwpos   : percentage F in that age with at least one adult female worm
Mnrw   : arithmetic mean nr of adult female worms per M in that age group
Fnrw   : arithmetic mean nr of adult female worms per F in that age group

The detailed *Y.txt output (e.g *example_STHY.txt*) contains sex and age specific output on skin snips with the first two columns indicating year and age group

year   : time (years)
age   : upper limit of age group

The remaining columns depend on the skin snip categories defined in the input file.

The default value of the `skin-snip.categories` attribute of the `<surveillance>` element is `skin-snip.categories="0.5,1,2,4,8,16,32,64,128,256,512,1e6"`

which results in the following output columns:
-1   : percentage of M in that age group with average skin snip count < 0.5
-1   : percentage of M in that age group with 0.5 <= average skin snip count < 1
-2   : percentage of M in that age group with 1 <= average skin snip count < 2
-4   : percentage of M in that age group with 2 <= average skin snip count < 4
-8   : percentage of M in that age group with 4 <= average skin snip count < 8
-16   : percentage of M in that age group with 8 <= average skin snip count < 16
-32   : percentage of M in that age group with 16 <= average skin snip count < 32
-64   : percentage of M in that age group with 32 <= average skin snip count < 64
-128   : percentage of M in that age group with 64 <= average skin snip count < 128
-256   : percentage of M in that age group with 128 <= average skin snip count < 256
-512   : percentage of M in that age group with 256 <= average skin snip count < 512
-1e9   : percentage of M in that age group with 512 <= average skin snip count < 1e9
followed by the same categories for F.

References

1. Coffeng LE, Stolk WA, Hoerauf A, Habbema D, Bakker R, et al. (2014) Elimination of African onchocerciasis: modeling the impact of increasing the frequency of ivermectin mass treatment. *PLoS One* **9**: e115886.

2. Plaisier AP, van Oortmarssen GJ, Habbema JD, Remme J, Alley ES (1990) ONCHOSIM: a model and computer simulation program for the transmission and control of onchocerciasis. *Comput Methods Programs Biomed* **31**: 43–56.

3. De Vlas SJ, Van Oortmarssen GJ, Gryseels B, Polderman AM, Plaisier AP, et al. (1996) SCHISTOSIM: a microsimulation model for the epidemiology and control of schistosomiasis. *Am J Trop Med Hyg* **55**: 170–175.

4. Plaisier AP, Subramanian S, Das PK, Souza W, Lapa T, et al. (1998) The LYMFASIM simulation program for modeling lymphatic filariasis and its control. *Method Inf Med* **37**: 97–108.

5. United Nations Department of Economic and Social Affairs Population Division (2013) World Population Prospects: The 2012 Revision, Volume I: Comprehensive Tables.

6. Montresor A, À Porta N, Albonico M, Gabrielli AF, Jankovic D, et al. (2015) Soil-transmitted helminthiasis: the relationship between prevalence and classes of intensity of infection. *Trans R Soc Trop Med Hyg* **109**: 262–267.

7. Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, et al. (2006) Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *Lancet* **367**: 1521–1532.

8. Anderson RM, Truscott J, Hollingsworth TD (2014) The coverage and frequency of mass drug administration required to eliminate persistent transmission of soil-transmitted helminths. *Philos Trans R Soc L B Biol Sci* **369**: 20130435.

9. Truscott JE, Hollingsworth TD, Brooker SJ, Anderson RM (2014) Can chemotherapy alone eliminate the transmission of soil transmitted helminths? *Parasit Vectors* **7**: 266.

10. Hotez PJ, Brooker S, Bethony JM, Bottazzi ME, Loukas A, et al. (2004) Hookworm infection. *N Engl J Med* **351**: 799–807.

11. Brooker S, Bethony J, Hotez PJ (2004) Human hookworm infection in the 21st century. *Adv Parasitol* **58**: 197–288.

12. Anderson RM, Schad GA (1985) Hookworm burdens and faecal egg counts: an analysis of the biological basis of variation. *Trans R Soc Trop Med Hyg* **79**: 812–825.

13. Montresor A, Urbani C, Camara B, Bha AB, Albonico M, et al. (1997) [Preliminary survey of a school health program implementation in Guinea]. *Med Trop* **57**: 294–298.

14.   Albonico M, Bickle Q, Ramsan M, Montresor A, Savioli L, et al. (2003) Efficacy of mebendazole and levamisole alone or in combination against intestinal nematode infections after repeated targeted mebendazole treatment in Zanzibar. *Bull World Heal Organ* **81**: 343–352.

15.   Gabrielli AF, Ramsan M, Naumann C, Tsogzolmaa D, Bojang B, et al. (2005) Soil-transmitted helminths and haemoglobin status among Afghan children in World Food Programme assisted schools. *J Helminthol* **79**: 381–384.

16.   Phommasack B, Saklokham K, Chanthavisouk C, Nakhonesid-Fish V, Strandgaard H, et al. (2008) Coverage and costs of a school deworming programme in 2007 targeting all primary schools in Lao PDR. *Trans R Soc Trop Med Hyg* **102**: 1201–1206.

17.   Pasricha S-R, Caruana SR, Phuc TQ, Casey GJ, Jolley D, et al. (2008) Anemia, iron deficiency, meat consumption, and hookworm infection in women of reproductive age in northwest Vietnam. *Am J Trop Med Hyg* **78**: 375–381.

18.   Vercruysse J, Behnke JM, Albonico M, Ame SM, Angebault C, et al. (2011) Assessment of the anthelmintic efficacy of albendazole in school children in seven countries where soil-transmitted helminths are endemic. *PLoS Negl Trop Dis* **5**: e948.

19.   Levecke B, Montresor A, Albonico M, Ame SM, Behnke JM, et al. (2014) Assessment of anthelmintic efficacy of mebendazole in school children in six countries where soil-transmitted helminths are endemic. *PLoS Negl Trop Dis* **8**: e3204.

20.   Augustine DL (1923) Investigations on the control of hookworm disease. XVI. Length of life of hookworm larvae from the stools of different individuals. *Am J Epidemiol* **3**: 127–136.

21.   World Health Organization (2012) Accelerating work to overcome the global impact of neglected tropical diseases - a roadmap for implementation.

22.   World Health Organization (2013) Sustaining the drive to overcome the global impact of neglected tropical diseases. Geneva.

23.   Habbema JDF, Oostmarssen GJ, Plaisier AP (1996) The ONCHOSIM model and its use in decision support for river blindness control. In: Isham V, Medley G. *Models for infectious human diseases - their stucture and relation to data*. Cambridge: Cambridge University Press. pp. 360–380.

24.   Pullan RL, Kabatereine NB, Quinnell RJ, Brooker S (2010) Spatial and Genetic Epidemiology of Hookworm in a Rural Community in Uganda. *PLoS Negl Trop Dis* **4**: e713.

25.   Anderson RM, May RM (1985) Herd immunity to helminth infection and implications for parasite control. *Nature* **315**: 493–496.